Available online at www.sciencedirect.com
SciVerse ScienceDirect
Physics Procedia 25 (2012) 485 – 491
Physics Procedia

2012 International Conference on Solid State Devices and Materials Science

The Research of Image Quality Assessment Methods

Xiaonan Cui[1], Zhiyuan Shi[1]*, Jianan Lin[1] , Lianfen Huang[2]

[1]*Dept. of Electronic Engineering, Xiamen University, Xiamen, China*
[2]*Dept. of Communication Engineering, Xiamen University, Xiamen, China*
**Corresponding author*

Abstract

In digital transmission, images may undergo quality degradation due to lossy compression and error-prone channels. Efficient measurement tools are needed to quantify induced distortions and to predict their impact on perceived quality. In this paper, an artiſcial neural network (ANN) is proposed for perceptual image quality assessment. The quality prediction is based on image features such as EPSNR, blocking, and blur. Training and testing of the ANN are performed with the mean opinion scores (MOS) provided by the Laboratory for Image and Video Engineering (LIVE). It is shown that the proposed image quality assessment model is capable of predicting MOS of the five types' image distortions.

Keywords: Image quality assessment; Artiſcial neural network; MOS; LIVE

1. Introduction

Measurement of the compressed image quality still remains an important issue in various images processing application, such as image acquisition, the choice of the parameters in coding system and performance comparison with some image compression algorithms.

Typically, the image quality assessment can be evaluated by subjective and objective evaluation. Subjective evaluation such as Mean Opinion Score (MOS), is truly definitive but too inconvenient, the most time taken and expensive [1] while objective evaluation is defined by mathematical definition, such as MSE, MAE, PSNR [2]. Generally, objective evaluation is based on the difference error of pixel values between two compared images and easy to calculate and usually has low computational complexity. However, they are not very well matched to visual quality perception [3]. So, more and more models have been provided.

In 2008, the International Telecommunication Union (ITU) publiced Recommendation J.247 which contains Yonsei University Full Reference Method. It is observed that the human visual system is sensitive to degradation around the edges. Based on this observation, the model provides a method that measures Edge Peak Signal to Noise Ratio(EPSNR). In this paper, a new model is applied according to the Yonsei Model with my own understanding. Further more an artiſcial neural network is adopted to compute MOS to make the prediction more accurate.

 doi:10.1016/j.phpro.2012.03.115

2. Preliminaries

2.1 LIVE Datebase

At Laboratory for Image and Video Engineering (LIVE), an extensive experiment was conducted to obtain scores from human subjects for a number of images distorted with different distortion types. These images were acquired in support of a research project on generic shape matching and recognition[4].

Twenty-nine high-resolution 24-bits/pixel RGB color images (typically 768 by 512) were distorted using five distortion types: JPEG2000, JPEG, white noise in the RGB components, Gaussian blur, and transmission errors in the JPEG2000 bit stream using a fast-fading Rayleigh channel model. A database was derived from the 29 images such that each image had test versions with each distortion type, and for each distortion type the perceptual quality roughly covered the entire quality range. Observers were asked to provide their perception of quality which is called mean opinion score (MOS) on a continuous linear scale that was divided into five equal regions marked with adjectives ``Bad", ``Poor", ``Fair", ``Good" and ``Excellent".

2.2 The Theory of Artificial Neural Network

An Artificial Neural Network (ANN) is an information processing paradigm that is inspired by the way biological nervous systems, such as the brain, process information[5].It is used to compute the MOS in this paper.

The commonest type of artificial neural network consists of three groups, or layers, of units: a layer of "input" units is connected to a layer of "hidden" units, which is connected to a layer of "output" units. The activity of the input units represents the raw information that is fed into the network. The activity of each hidden unit is determined by the activities of the input units and the weights on the connections between the input and the hidden units. The behaviour of the output units depends on the activity of the hidden units and the weights between the hidden and output units[6].

The linear network is made up of many linear neurons, shown at the following figures.

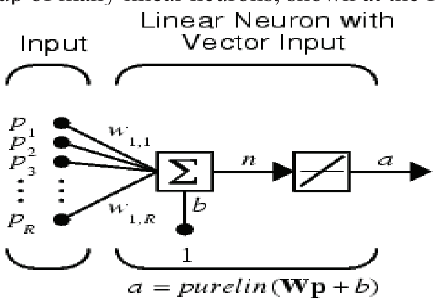

Figure 1. A linear neuron model.

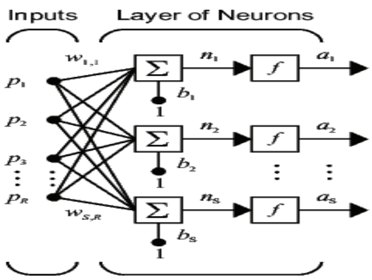

Figure 2. A linear network model.

Linear network use the LMS algorithm or Widrow-Hoff learning algorithm based on an approximate steepest descent procedure[7]. Surpose there is a training base with m samples. {p1，t1}{p2, T2}….. {PM, TM}, pm isinput, tm is target output， am is network output, and error is e (k) =t (k)-a (k).

$$mse = \frac{1}{m}\sum_{k=1}^{m} e^2(k) = \frac{1}{m}\sum_{k=1}^{m}\left(t(k) - a(k)\right)^2 \tag{2.1}$$

The LMS algorithm adjusts the weights and biases of the network so as to minimize this mean square error. The details of the rule can be found as follow:
First, initialize the weights and baise with a small random value.
Second, input a sample to compute the adjustment.

$$\frac{\partial e(k)}{\partial w_{ij}} = \frac{\partial e}{\partial w_{ij}}\left[d(k) - \left(\sum_{i=1}^{R} w_{ij} p_i(k) + b\right)\right] \quad j = 1, 2, \cdots S \tag{2.2}$$

$$\frac{\partial e^2(k)}{\partial w_{ij}} = 2e(k)\frac{\partial e(k)}{\partial w_{ij}} \tag{2.3}$$

$$\frac{\partial e^2(k)}{\partial b} = 2e(k)\frac{\partial e(k)}{\partial b} \tag{2.4}$$

Here $p_i(k)$ is the ith element of the input vector at the kth iteration.so:

$$\frac{\partial e(k)}{\partial w_{ij}} = -p_i(k) \tag{2.5}$$

$$\frac{\partial e(k)}{\partial b} = -1 \tag{2.6}$$

Third, change the weights and the bias as they shown below

$$w(k+1) = w(k) + 2\eta e(k) p^{\mathrm{T}}(k) \tag{2.7}$$

$$b(k+1) = b(k) + 2\eta e(k) \tag{2.8}$$

Here η is learning rate. To ensure stable learning, the learning rate must be less than the reciprocal of the largest eigenvalue of the correlation matrix $p^T p$ of the input vectors.

Finaly, calculate the error using the equation 2.1, if the error reaches the goal the algorithm will end, else go back to the second step to have a new cycle.

3. The Model Of Image Quality Assessment

After understanding the LIVE database and artificial neural network model, an image quality assessment model can be developed. It could be realised by two steps: feature extraction and MOS calculation.

3.1 EPSNR Module

1) Feature Extraction

In this module, EPSNR (Edge Peak Signal to Noise Ratio) blocking and blurriness features are treated as features to compute MOS.

a) EPSNR

First, a sobel edge detection algorithm is applied to compute the horizontal gradient image $g_{horizontal}(m,n)$ and the vertical gradient image $g_{vertical}(m,n)$. Then, the magnitude gradient image g(m, n) may be computed as follows:

$$g(m,n) = | g_{horizontal}(m,n) | + | g_{vertical}(m,n) | \tag{3.1}$$

Second, a thresholding operation is used to the magnitude gradient image g(m, n) to find edge pixels and its quantity.

Finally, the EPSNR of the source and processed images can be computed as follows:

$$MSE_{sdge} = \frac{1}{K}\sum_{i=1}^{M}\sum_{j=1}^{N}\left\{S(i,j)-O(i,j)\right\}^2. \tag{3.2}$$

$$EPSNR = 10\log_{10}(\frac{p^2}{MSE_{edge}}) \tag{3.3}$$

where p is the peak value of the image.

b) blocking and blurriness features

In order to extract features which measure the degrees of blocking and blurriness of the processed images, the model first extracts edge pixels and computes horizontal (H (i, j)) and vertical (V (i, j)) gradient component of the edge pixels with Sobel operators. From the horizontal and vertical gradient images, the magnitude (R (i, j)) and angle (θ (i, j)) are computed as follows:

$$\begin{aligned} R(i,j) &= \sqrt{H(i,j)^2 + V(i,j)^2} \\ \theta(i,j) &= \tan^{-1}\left[\frac{V(i,j)}{H(i,j)}\right] \end{aligned} \tag{3.4}$$

Then, the horizon and vertical component (HV (i, j)) is computed as follows:

$$HV(i,j) = \begin{cases} R(i,j), & R(i,j) \geq r_{\min} and\, m\frac{\pi}{2} - \Delta\theta < \theta(j) < m\frac{\pi}{2} + \Delta\theta, \Delta\theta = 0.225 \\ 0 & \end{cases} \tag{3.5}$$

Finally, the blocking feature (Fblocking) is computed as follows:

$$F_{blocking} = \frac{1}{n_{blocking}}\sum_k (HVp(k) - HVs(k)), if\,(HVp(k) > HVs(k)) \tag{3.6}$$

From the source images, the model produces a number of HV (i, j), which is denoted as {HVs (k)}. It is noted that all pixels which satisfy the condition ($r_{min} \geqslant 110$) are used in this procedure. From the processed images, the model generates a number of HV (i, j) , which is denoted as { HVp(k) }.where nblocking is the number of pixels satisfying the condition (HVp(k) > HVs(k)) .

Furthermore, the blurriness feature (Fblur) is computed as follows:

$$F_{blur} = \frac{1}{n_{blur}}\sum_k (HVs(k) - HVp(k)), if\,(HVs(k) > HVp(k)) \tag{3.7}$$

2) MOS Calculation.

Above, the overall aim is to design an ANN that can assess and quantify image quality in terms of predicted MOS. Accordingly, the favorable ANN needs to be trained to ﬁnd associations between input signals (image features) and a corresponding desired response (predicted MOS). Clearly, the trained neural network should not only be able to map known inputs to known outputs but should also be able to associate unknown inputs to meaningful outputs. In the sequel, we will present the considered linear network architecture and describe its training and testing[8].

The MOS can be computed by a linear combination of the three features (EPSNR, F_blocking, and F_blur) as follows:

$$\Pr\_MOS = a \times EPSNR + b \times F\_blocking + c \times F\_blur \tag{3.8}$$

Known from the above formula, the network has three inputs called EPSNR, F_blocking and F_blur,and a, b, c are their respective weights. Pr_MOS is the output.

In order to establish a linear neural network whice has good generalization ability, sample selection is extremely important. In this paper, the samples consist of all the five distortion types, and are selected from every range of the features.

In this paper, 223 samples have been set to train the network..After training and testing the network, a = 0.093, b = 0.0041, c =- 0.0012, and RMSE = 0.4100.

The figure below shows the results. The horizontal axis indicates the predicted MOS, the vertical axis means the MOS in the LIVE datebase, and the dots stands for datas.

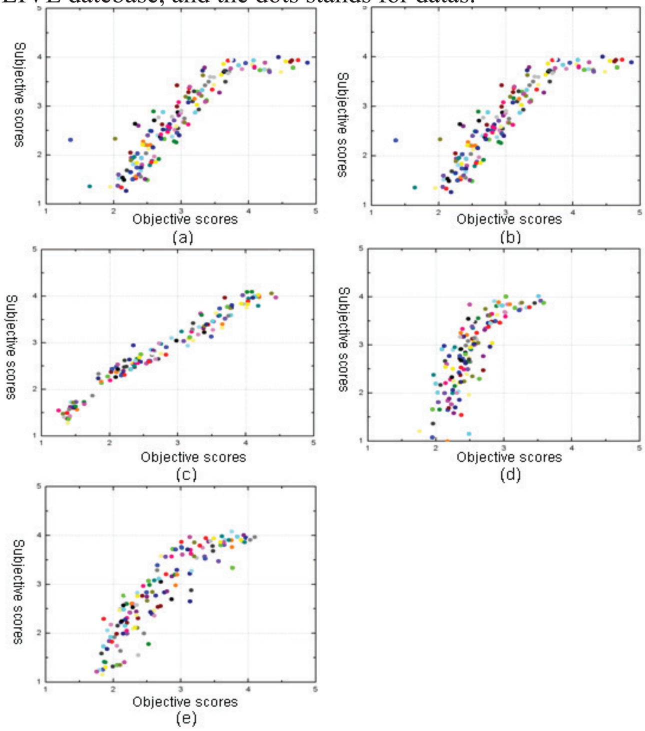

Figure 3. Comparison of the subject and object scores for EPSNR model. Picture(a) is the comparison for JPEG 2000 images, Picture (b) is for JPEG images,picture (c) is for white noise images,picture (d) is for Gaussian blur images,and picture (e) is for fast_fading images..

3.2 Modified EPSNR module—MEPSNR module

For most of the JPEG, JPEG2000 and fast Rayleigh fading image, the evaluation results are good, but in the higher part of the MOS the results turn up a larger deviation. The computed MOS are higher then the normal MOS.To addresses this issue, this paper propounds a nonlinear corrected on EPSNR called MEPSNR Model.Specific improvements are as follows:

$$MEPSNR = \begin{cases} EPSNR & if \ 0 \le EPSNR \le 35 \\ EPSNR \times 0.9 & if \ 35 < EPSNR \le 40 \\ EPSNR \times 0.8 & if \quad EPSNR > 40 \end{cases} \tag{3.9}$$

$$MEPSNR = \begin{cases} MEPSNR - 60 \times \left[0.1225 - \left(\frac{EP_c}{EP_s} \right)^2 \right] & if \quad EPSNR < 25 \& \frac{EP_c}{EP_s} < 0.35 \& \frac{EP_p}{EP_s} < 0.13 \\ MEPSNR \quad elsewhere \end{cases} \tag{3.10}$$

EP_s is the edge pixels quantity of the source images, EP_p is the edge pixels quantity of the processed images, EP_c is the edge pixels quantity of both. And the blocking and blurriness features stay the same.

$$\Pr\_MOS = a \times MEPSNR + b \times F\_blocking + c \times F\_blur \tag{3.11}$$

The linear neural network model is also used here to compute MOS. The first input of the network is replaced by MEPSNR, but others remained. After training and testing the network, a=0.1026 , b=0.0032 , c=-0.0012, and RMSE=0.3954.

The reasults are shown below.

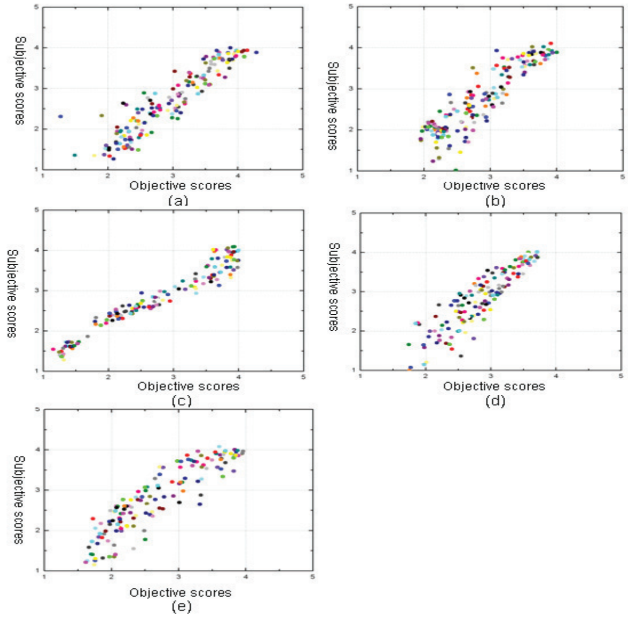

Figure 4. Comparison of the subject and object scores for MEPSNR model Picture(a) is the comparison for JPEG 2000 images, picture (b) is for JPEG images, picture (c) is for white noise images, picture (d) is for Gaussian blur images, and picture (e) is for fast_fading images.

4. Experimental Results and Analysis

The experimental results are listed in the table:

TABLE I. Experimental results

Images	**performance parameter**	**EPSNR**	**MEPSNR**
JPEG2000	***RMSE***	0.4538	0.4200
	Pearson correlation	0.8992	0.9285
	Outlier Ratio	0.0473	0.0236
JPEG	***RMSE***	0.4311	0.4063
	Pearson correlation	0.8859	0.9080
	Outlier Ratio	0.0342	0.0228
White noise	***RMSE***	0.1695	0.1957
	Pearson correlation	0.9818	0.9744
Gaussian blur	***RMSE***	0.5051	0.5126
	Pearson correlation	0.8194	0.8070
Fast-Fading Rayleigh channel	***RMSE***	0.3922	0.3649
	Pearson correlation	0.8961	0.9071

This paper established two assessment models: EPSNR and MEPSNR. Among them MEPSNR is the modified module. Improved results can be seen from Figure 4.(a), the higher part of the MOS has had a good correction, and its value is uniformly distributed near the diagonal. The two models also have great performance, which can be found from Table I.

5. Conclusion

Now, two models have been established for image quality assessment. As a result, the method of establishing assessment model can be summarized as follows:

Step 1: Prepare subjective database. In order to get an effective model a reliably and authorized database is required. The datas should be highly reliable; images must be sufficient, far-ranging and representative In this paper, an overt LIVE subjective database which is widely used and approved is adopted.

Step 2: Set features. Features must change effectively according to the image quality varies. The features in this paper are more mature.

Step 3: Compute these features for the next stage.

Step 4: Predict human's subjective feelings with the features. Compare the predicted results to the nomal MOS and compute the MOS with a certain criteria, such as the minimum criteria for RMSE used in this paper. When calculating MOS, the artificial neural network technology has been integrated well and a linear network has been established.

Step 5: Evaluate the model's performance completely. It is important to evaluate the model's performance with other aspects, such as correlation coefficient, outlier ratio and so on to summarize the advantages and disadvantages of the model.

Acknowlegment

The authors would like to thank the Cooperation Project of Xiamen University and Xiamen Municipal Science and Technology Agency for their financial support.

References

[1] ITU, "Methods for the Subjective Assessment of the Quality of Television Picture," ITU-R Rec. BT. 500-7, 1998.

[2] M. Eskicioglu and P.S. Fisher, "Image Quality Measures and Their Performance," IEEE Transactions on Communications, vol. 43, no. 12, 1995, pp. 2959-2965.

[3] S. Grgic, M. Grgic, and M. Mrak, "Reliability of Objective Picture Quality Measurement," Journal of Electrical Engineering, vol. 55, no. 1-2, 2004, pp. 3-10.

[4] H. R. Sheikh, Z. Wang, L. Cormack, and A. C. Bovik, LIVE ImageQuality Assessment Database, Release 2 2005 [Online]. Available: http://live.ece.utexas.edu/research/quality

[5] P. Carrai, I. Heynderich, P. Gstaldo, and R Zunino, "Image quality assessment by using neural networks," Pmc. IEEE Inter. Symposium on Circuits and Systems (ISCAS-2001), pp. V-253-256, 6-9 May 200 I, Sydney.

[6] J. M. Zurada, introduction to urfiicial neural system: PWS publishiner company, 1992. S. Daly. The visible difference predictor: an algorithm for the assessment of image fidelity. In A. B. Watson, editor, Digital Images and Human Vision, pages 179- 206. MIT Press, Cambridge, Massachusetts, 1993..

[7] E. Oja. Neural networks, principal components and subspace. International Journal of Neural Systems, 1:61-68, 1989.

[8] Zhou Wang, Guixing Wu, H.R. Sheikh, E.P. Simoncelli, En-Hui Yang, A.C. Bovik, "Quality Aware Images," IEEE Transactions On Image Processing, Vol. 15, No. 6, June 2006, pp. 1680-1689